AFRL-OSR-VA-TR-2014-0342

MULTI-LEVEL CULTURAL MODELS

Kathleen Carley
CARNEGIE MELLON UNIVERSITY

11/05/2014
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTC
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
29-08-2014	Final Report	01-10-2011 - 31-5-2014

4. TITLE AND SUBTITLE
Multi-Level Cultural Models

5a. CONTRACT NUMBER
E2016762

5b. GRANT NUMBER
FA9550-11-1-0179

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kathleen M. Carley

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Ben Knott
Air Force Office of Scientific Research
801 N Randolph St.
Arlington VA 22203
United States

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A - Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Empirically grounded multi-agent meta-network multi-level simulations were developed, tested, validated and then used to explain and predict socio-cultural outcomes. Virtual experiments employing real and virtual data were conducted examining the impact of group structure and the communication media on the evolution of the group, the culture and the change in sentiment. The Arab Spring was used as a case example. A formal theory of socio-cultural dynamics based on the co-evolution of self and the groups one is a member of was instantiated as multi-level construct. The results indicate that: model re-use for social-network diffusion models is feasible; adding social cognition to a multi-level model improves realism, accuracy of predictions, and decreases memory consumption and improves run-time speed.

15. SUBJECT TERMS

dynamic network analysis, ABM, agent based models, simulation, re-use, text mining, social networks, open source exploitation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Kathleen M. Carley
unclass	unclass	unclass			**19b. TELEPHONE NUMBER** *(include area code)* 412-268-6016

Multi-Level Cultural Models

Final Report

FA8750-08-2-0020 E2016762

Kathleen M. Carley

Today's Air Force engages in missions requiring it to work with and understand other organizations or militaries which may have vastly different cultures; e.g., deterrence, humanitarian relief, peacekeeping operations, and multi-national conflict. This creates a need to understand the movement of ideas and beliefs and to assess resiliency. Socio-cultural networks within and among organizations influence behavior, constrain information diffusion, and impact the system's resiliency and the level of trust. Cultural-forms, such as hierarchy, constrain these and create expectations in agents about other's behavior. Individual's perception of the generalized other (i.e., "what-every-one-knows" and "what-every-one-does") impacts their willingness to engage in diverse activities from providing aid to those hit by disasters to harboring terrorists to engaging in acts of violence. The situation is further complicated by the fact that migration, wars, natural disasters, global warming and changing economic conditions often result in massive population shifts at an unprecedented rate thus altering cultures and creating instability in the underlying beliefs, norms and behaviors.

Socio-cultural networks within and among organizations influence behavior, constrain information diffusion, and impact the system's resiliency and level of trust. These socio-cultural networks link people, groups, ideas and beliefs in a complex web of relations that changes through time and space. If we are to explain and predict socio-cultural outcomes, such as those between competing or collaborating factions within geographic regions, we need to be able to explain and predict these socio-cultural networks. Our approach was to use semi-auto-instantiated multi-level agent-based dynamic-network models to explore social movements toward or away from behavioral outcomes of interest as individual agents in the model engage in social interactions. A high level view is shown in Figure 1.

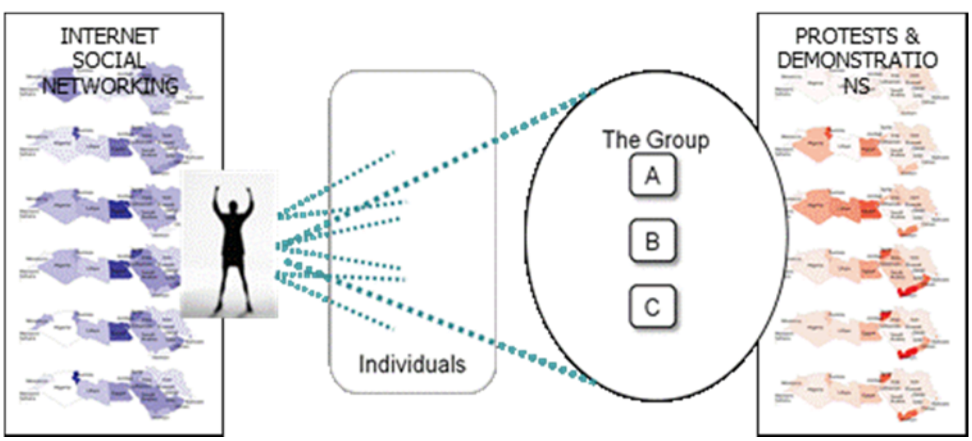

Figure 1. Newspaper data on the Arab Spring is used to instantiate a model in which individuals interact with each other, join and leave groups, and change in their perception of the logic of engaging in violent acts or protests. Interactions are direct or through various media.

The research described herein employed empirically grounded multi-agent meta-network multi-level simulations to explain and predict socio-cultural outcomes. Virtual experiments employing real and virtual data were conducted focusing around three key questions: First, how

do beliefs, norms, and behavior within and among groups or factions change as populations shift? Second, how do preferences for specific cultural forms and perception of the generalized other interact with the structure of the social network, new members, and the changing demographics, to effect change, promote or inhibit adversarial relations, and enable the emergence of new cultures and cultural forms? Third, are there means of communicating (such as use of social media versus face-to-face communication) with the members of a group or faction that are more effective at maintaining the current culture or effecting cultural transformations, and so altering tendencies toward various behaviors such as providing social support or engaging in violence. The first objective was to develop a formal theory of socio-cultural dynamics based on the co-evolution of self and the groups one is a member of. The second objective was to create, test and validate a formal model of cultural forms that accounts for population shifts, perception and assessment of the generalized other, and the movement of individuals into and out of groups in response to changing socio-cultural conditions. This theory was to be instantiated as a multi-level agent-based dynamic-network simulation. The third objective was to develop procedures for auto-instantiating and creating re-usable simulations. This procedure was to be tested using real-world data collected from open-source venues. The final system would support rapid sense-making given auto-instantiated dynamic network agent based models. This in turn provided the foundational research that enables the AirForce to understand, predict, and influence the flow of information, beliefs, and associated cultural shifts, and to identify impact of mode of communicating with US, adversarial and allied groups and populations.

Data

The data used were networks extracted from English newspaper stories for 16 countries in the Middle East over a period of time encompassing the Arab Spring. The countries are shown in Figure 2. A total of 400,000 items were scraped from Lexis-Nexis covering 16 countries for 20 months from August 2010 through March 2012. This data was segmented by country by month.

Figure 2. Data for each country shown was collected from Lexis-Nexis.

Using a set of text-mining scripts, the Stanford entity extractor, and AutoMap these articles were mined to produce a set of networks of actors and topics, by country by month. These topics were then cross-categorized by violence and protest to determine which were pro/con violence and

which pro/con protest. A mask linking the beliefs pro/con violence and pro/con protest was then created such that a topic had a positive one link if it was pro, a negative one link if it was con and a 0 link if it was neutral. Each actor was linked to just those topics that they were co-mentioned with in the articles within a 2 sentence window. All actors were vetted and those known to be US citizens and not major international actors were removed. Major international actors – e.g., the US president, US vice president, US secretary of State, were kept in the data set. Based on this information the basic pro/con level of each actor for each of violence and protest was estimated. For the top political actors, i.e. leader of each country and major US international actors, if they were neutral on these beliefs the following beliefs were assigned to con-violence and con-protest. This data was used to auto-instantiate the model in experiment 3. Finally, for each country a generalized other was added to represent the general public and this was linked to the average of the real data.

Using web-based reviews of the Arab Spring, a check list of basic outcomes: when first protest occurred, when leader was over-turned, when violence abated was developed. This was used for checking validity of results. In addition, part of the news data was used to estimate level of violence and protest. This too was used to check validity of results.

Multi-level Construct Model

A multi-level version of Construct was implemented, tested, validated, and used to assess social change. Construct is an agent-based dynamic-network simulation system design to allow the user to assess the spread of information and beliefs/sentiments across populations in which the underlying social networks and culture co-evolve. In Construct agents can communicate via different media at the simulators choice. Options include face-to-face communication, group/public lectures, email, newspapers, phone, twitter, and so on.

The original single level Construct assumed that all actors were essentially similar and acted as though they were at the same ontological level. In fact, humans, operate at both an individual and a group level and use social intelligence, their knowledge about groups, to infer information, to categorize information, and to make decisions. We implemented a multi-level version of Construct by adding social intelligence. A high-level view of multi-level Construct is shown in Figure 3.

There are several features that make Construct now multi-level. First the actors can be both single actors or meta-actors (collections of actors). Second individual actors know whether they are dealing with an single or a meta-actor and reason accordingly. Actors treat meta-actors as a "generalized other" which has the "knowledge associated with it" of the average member. When actors meet a new actor they infer what that new acquaintance will know based on group (meta-actor) membership. Actors form their view of the group based on those they have met. Actors in active memory store and retrieve data for those that they are not strongly tied to vis the meta-actor information.

In this multi-level version it is the case that the user can make use of:

- Multiple level of actors – meta-actors (e.g. groups contain other actors) & specific actors
- Knowledge/Perceptions/Beliefs that are multi-level
  - Knowledge/perceptions/beliefs about "the Group", the Generalized Other (Mead)
  - Knowledge/perceptions/beliefs about individuals within that group, the Specific Others
- Actors make decisions at multi-levels and actors at different levels make actions at different levels

- Actions depend on the level of knowledge/perceptions/beliefs, such that:
  - Actors interact with members of groups based on generalized info about the group (stereotyping)
  - Actors learn/gain both specific and general knowledge/perceptions/beliefs
  - An actor's interaction with specific other actors forms the actor's group perception (generalization)

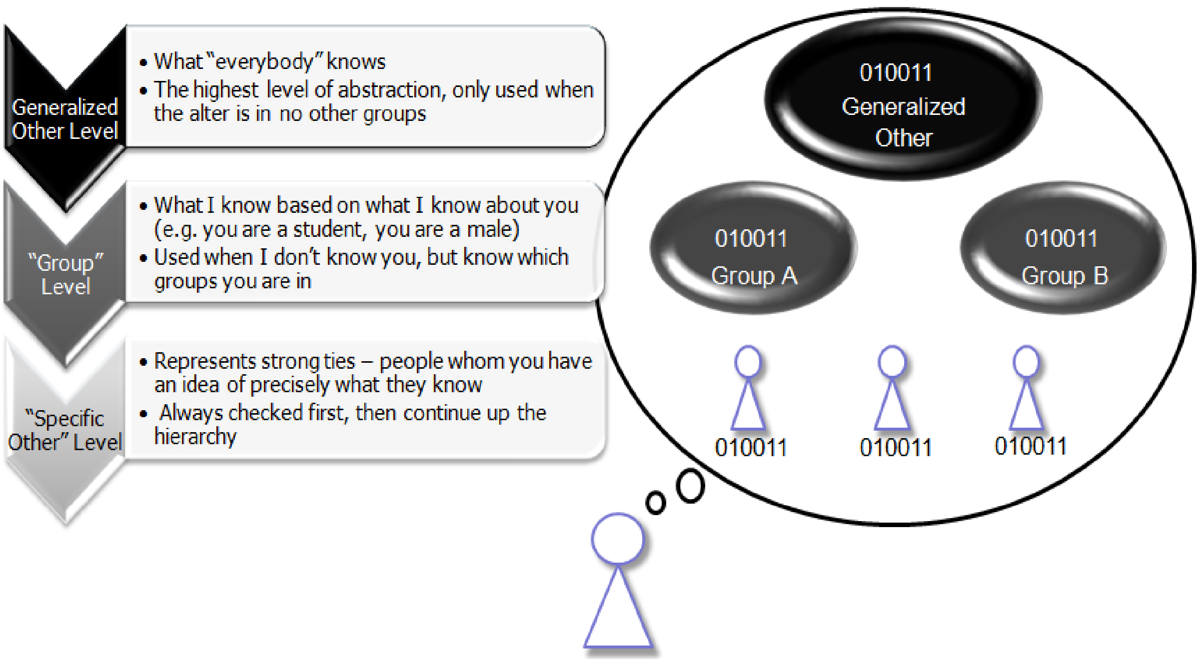

Figure 3. High level conceptualization of key features in multi-level Construct.

Thus in multi-level Construct there are:

- N levels of actors
  - Meta-actors
    - Statistical actors
    - have "preferences" and "knowledge" formed though lossy intersection of sub-groups
  - Sub-actors treat unknown others as generalized other using the most meta-group encountered
    - Agent's transactive memory of specific other is set by default to the meta-actor's
  - Joining
    - enables learning meta-actors actual structure and beliefs and alters actual beliefs
    - Increases interaction with associated specific actors
  - Leaving increases likelihood of forgetting meta-actors
- Actors
  - Are information processors in social networks with social-cognition

- Preferences, knowledge, sphere of influence
  - Includes knowledge about generalized others (meta-actors) and specific actors (transactive memory)
- Each actor has a cultural form
  - Associated internal interaction structure and beliefs which impacts level of consensus and rate of making decision

- Networks connect actors
  - Are observable through communications vis different media

Virtual Experiments

To conduct this research a set of virtual experiments were run using the original (single level) and the new multi-level Construct model. These experiments were designed for various purposes including to test, verify, and validate the model and to use the model to address various theoretical issues. Over the course of the project 20 major virtual experiments and a number of small virtual experiments were run. Only three selected experiments are described herein.

Key Results

Experiment 1: Impact on memory and speed. A virtual experiment was conducted using notional groups and actors to test the impact of adding social cognition to Construct on memory usage and time. In this experiment there were 1000 agents, with a realistic density (on the high side) of .125, less knowledge bits than agents (500 "facts") and a .1 density in the agent-knowledge network (each agent knows 10 percent of the knowledge. This was simulated for 50 time periods. The results are shown in Figure 4 (memory) and Figure 5 (time). The results show that multi-level Construct is linear in memory (thus, it is dependent only on the initial configuration and not on changes in density or network structure) and it is significantly faster. Memory scales as NM (number of agents x number of knowledge bits).

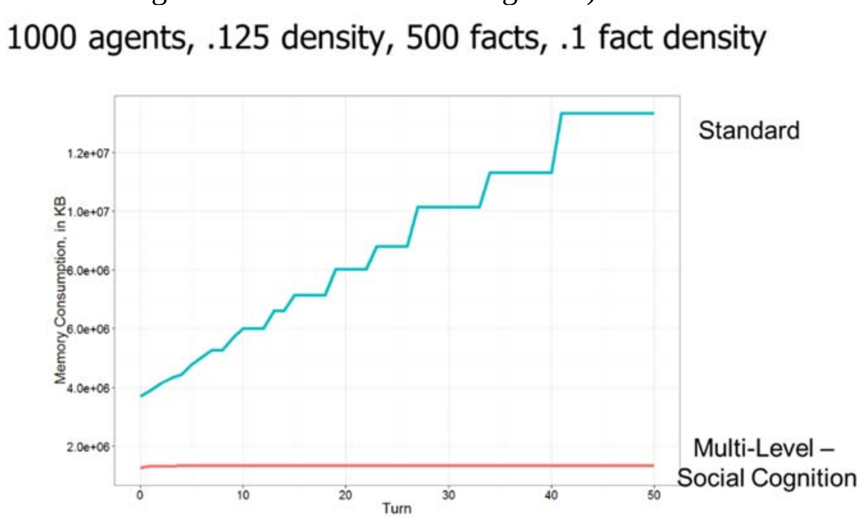

Figure 4. Change in memory usage over the course of a simulation experiment.

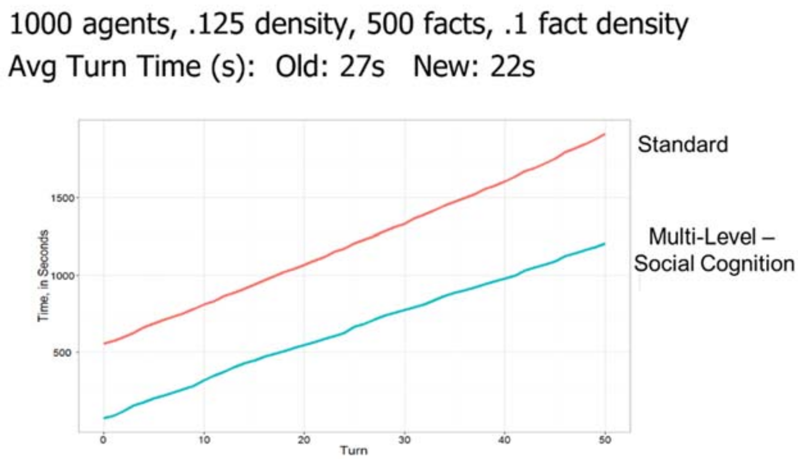

Figure 5. Change in real-time it takes to complete a set of turns (simulation communication and learning cycles).

Experiment 2: Structure and Cohesion. The goal was to conduct a theoretical examination of how structure and cohesion would evolve naturally over time in communities. In this virtual experiment the populations are stylized communities of actors that vary in their level of initial expertise, such that the number of groups and the number of actors was varied randomly. A set of 60,000 runs were conducted. The images shown for the results contain a random sample of 12,000 of these runs. The results of this study show that expertise is a driver of community behavior. In Figure 6, we see that there is an interaction between group size and the presence of ties between groups (cross-group ties), such that large groups need cross-group ties in order to achieve high performance. Over time, large groups forming cross-group ties outperform those that do not. Such a dynamic is not as needed in smaller groups for high performance. In Figure 7 the impact of ethnocentrism is shown. The new model has the feature that shared culture, social relations and group stereotypes all intermingle to produce macro-social structure. The results show that the model predicts that the more specialized the knowledge in a group, the less likely the members of that group are to form ties with those outside their group (left) and schemas about what other groups are like (right). Or to put it another way, based on this model the theory suggests that low ethno-centrism promotes group formation and out-group ties.

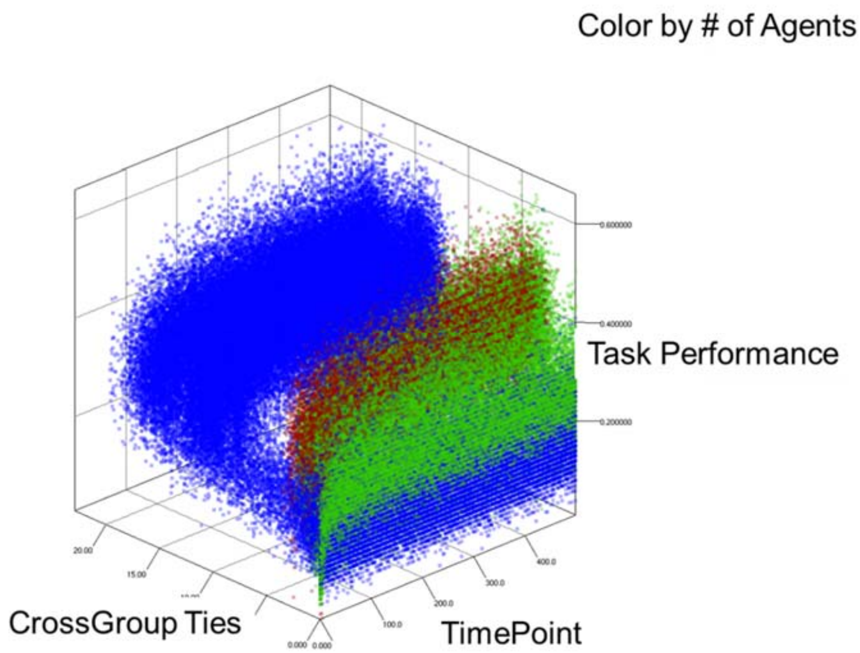

Figure 6. Expected performance as a function of cross-group ties and group-size.

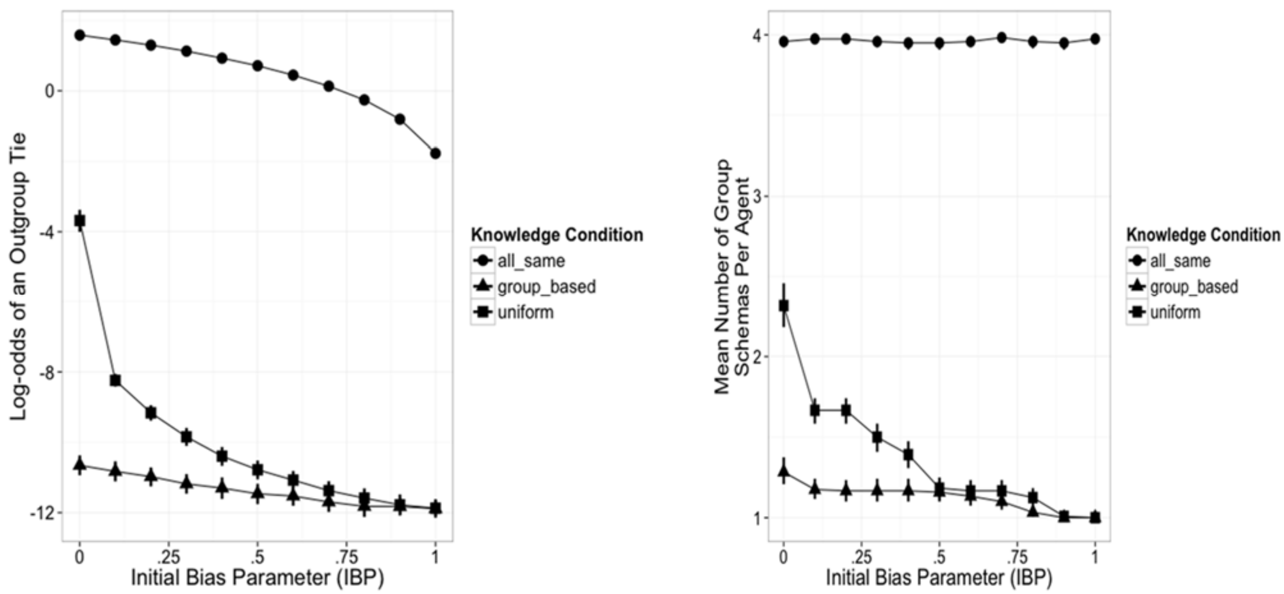

Figure 7.

Experiment 3. This was the Arab Spring Experiment. The goal was to see to what extent we could predict the change in sentiment and so the potential for outbreaks of violence and protest given the prior months data. The aim was rapid re-usable modeling. In this case each country was simulated month by month to predict the sentiment toward violence and protest in the following month. The model was instantiated with the data previously described. Eight replications were performed for each month of data, each with slightly different parameterizations of MLC. With eight replications of each month this is a total of 160 simulation runs, each of which simulates between 7000 and 13000 agents. The results are showing in Figure 8, with a dot showing the first outbreak of violent protest.

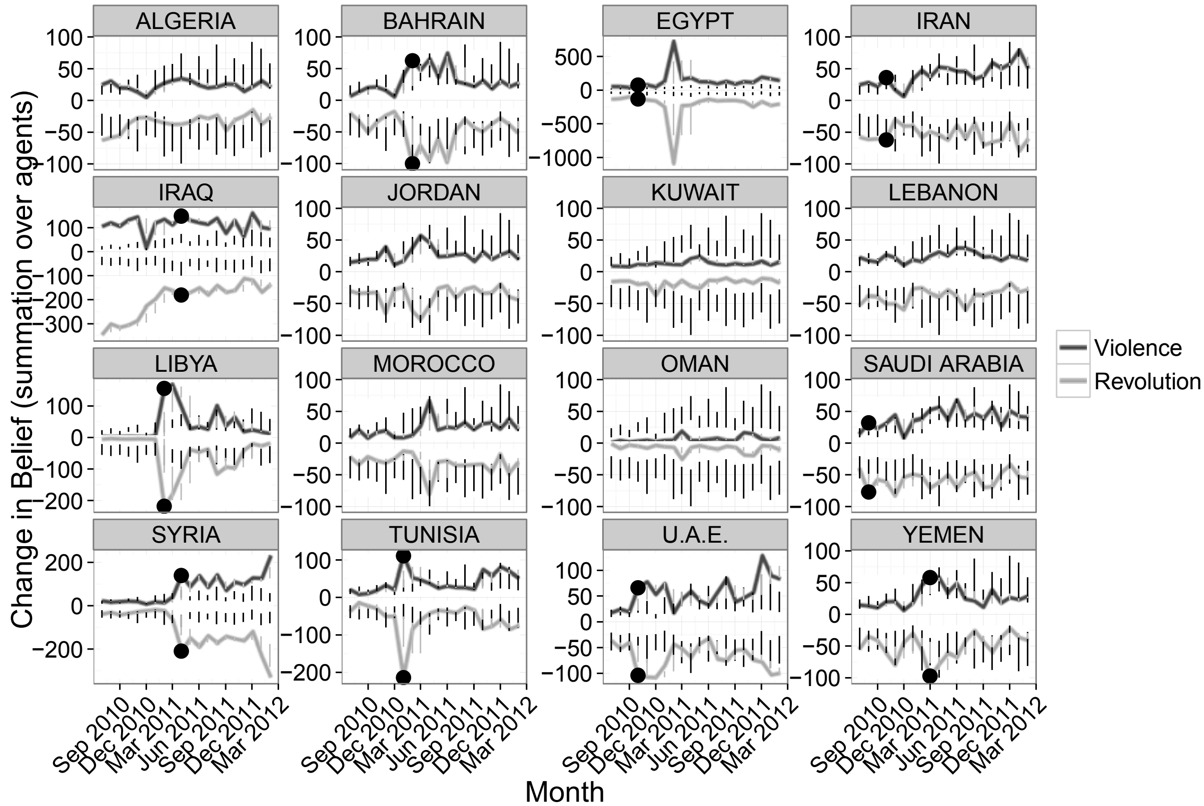

Figure 8. Change in population level sentiment pro/con violence and revolution for 16 countries by month as predicted by simulator. Dots indicate first outbreak of protest or violence.

Technical Challenges Addressed

Re-use:

Most agent-based simulation (ABM) in use today are one-off. In contrast, we demonstrated re-use through the identification of semi-automatically instantiable model components including census data (year 1) and text-mining data (years 2 and 3). Such an approach should work for any information diffusion or sentiment change model.

Multi-Level Modeling:

Most ABM have all agents at the same level; e.g., all people, or all groups. In contrast, we developed a true multi-level model using social-cognition in which there are actors at different levels and which were able to reason about those levels. We demonstrated that the temporal patterns of behavior in a multi-level system are the same as with a single level system. However, multi-leveling increases speed and computational efficiency of the model. Further, multi-level models show group formation as a naturally emergent behavior. And finally, the predictive accuracy of multi-level models for the same phenomena is higher when multi-level modeling is used.

Compatibility:

Typically simulation models are “designed” relative to a context. As such, they can only be used in that context. In contrast, we used a hierarchical modeling approach in which there was a modeling framework which was then instantiated for a general context which was then instantiated

for specific contexts. The modeling framework can be used across contexts. The general model can be used across specific contexts. And the specific models can be used across time for the same specific context. This hierarchical approach supported re-use. This was demonstrated in the Arab Spring experiments where we used the Multi-Level Construct (which has since been re-used in other contexts such as for information leaks and cyber security) as the framework. We then created a general protest model in which there are two beliefs violence and protest, a generalized other for the general public, individual actors and public leaders. We then built for each country a specific model with that country's actors. The raw data was used to re-instantiate each of these 20 times once per month.

Veridicality-Speed Tradeoff:

As the veridicality of an ABM increases (either more realistic social networks or more realistic cognition) the speed of processing decreases and number of agents modeled at the same speed decreases. This decrease is approximately linear in speed and approximately N^2 in the number of agents. We demonstrated that adding social cognition in a multi-level model enables increased speed, decreases memory and so enables more veridical larger populations to be modeled.

Access

The AutoMap system used for text-mining is available at: http://www.casos.cs.cmu.edu/projects/automap/ .

The Construct system used for the simulations is available at: http://www.casos.cs.cmu.edu/projects/construct/ .

Metrics

Education:

5 Ph.D. students were supported in whole or in part. The work herein led to three thesis proposals, and a new chapter in an on-going and now finished thesis.

All students attended and taught in the CASOS summer institute for all three years of the project.

The multi-level model and text to model process was taught in, and used by the participants in, the CASOS summer institute. The participants, roughly 30 to 50 per year, are a collection of PhD students, Faculty, Industry personnel, and Government personnel.

The multi-level model and text to model process was also taught in a new module in Dr. Carley's Masters/PhD course in simulation.

Publications - Journals:

Joseph, Kenneth, Kathleen M. Carley, David Filonuk, Geoffrey P. Morgan, and Jürgen Pfeffer, to appear. Arab Spring: From News Data to Forecasting. Social Networks and Mining.

Lanham, Michael & Morgan, Geoffrey & Carley, Kathleen M. (2014). Social Network Modeling and Agent-Based Simulation in Support of Crisis De-escalation. *IEEE: Transactions on Human-Machine Systems*, 44(1), 103-140. http://dx.doi.org/10.1109/TSMCC.2012.2230255

Joseph, Kenneth, Geoffrey P. Morgan, Michael K. Martin and Kathleen M. Carley. (2014). On the Coevolution of Stereotype, Culture, and Social Relationships: An Agent-Based Model. *Social Science Computer Review*. 32(295), Sage. 10.1177/0894439313511388

Publications – Theses:

Frantz, Terrill. (2013). A Behavioral Theory of the Merger. *Carnegie Mellon University, School of Computer Science, Institute for Software Research, Computation, Organizations and Society, Doctor of Philosophy.*

Publications – Conference Proceedings and Book Chapters:

Carley, Kathleen M & Morgan, Geoffrey & Lanham, Michael & Jürgen Pfeffer. (2012). Multi-Modeling and Sociocultural Complexity. *In: D.D. Schmorrow, D.M. Nicholson (Eds,), Advances in Design for Cross-Cultural Activities Part II, CRC Press, pp. 128-137.*

Carley, Kathleen M & Pfeffer, Jürgen. (2012). Dynamic Network Analysis (DNA) and ORA. *In: D. D. Schmorrow, D.M. Nicholson (Eds.), Advances in Design for Cross-Cultural Activities Part I, CRC Press, pp. 265-274.*

Wei, Wei & Carley, Kathleen M. (2014). Real Time Closeness and Betweenness Centrality Calculations on Streaming Network Data. *In the proceedings 2014 ASE BIGDATA/SOCIALCOM/CYBERSECURITY Conference, Stanford University, May 27-31, 2014*

Carley, Kathleen M & Morgan, Geoffrey & Lanham, Michael & Pfeffer, Jürgen. (2012). Multi-Modeling and Socio-cultural complexity: Reuse and Validation. *In proceedings of the 2nd International Conference on Cross-Cultural Decision Making: Focus 2012, San Francisco, CA, July 21-25, 2012.*

Carley, Kathleen M & Pfeffer, Jürgen. (2012). Dynamic Network Analysis (DNA) and ORA. *In proceedings of the 2nd International Conference on Cross-Cultural Decision Making: Focus 2012, San Francisco, CA, July 21-25, 2012.*

Joseph, Kenneth & Carley, Kathleen M. (2012). Quantifying Meaning in Interaction on Different Communication Media. *INSNA Sunbelt Conference, March 15, 2012.*

Carley, Kathleen M & Bigrigg, Michael & Garlan, David & Lanham, Michael & Lu, Yue & Morgan, Geoffrey & Schmerl, Bradley. (2011). Experimentation Testbeds: Using SORASCS to Run and Process HSCB Virtual Experiments. *In proceedings of HSCB Focus 2011: Integrating Social Science Theory and Analytic Methods for Operational Use, Chantilly, VA, February 8-10, 2011.*

Technical Reports

Carley, Kathleen M & Columbus, Dave & Landwehr, Peter. (2013). AutoMap User's Guide 2013. *Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISR-13-105*

Posters and Presentations:

2 Poster presentations at the CASOS Institute in 2012
3 Poster presentations at CASOS Institute in 2013
2 Poster presentations at the CASOS Institute in 2014

7 conference presentations
12 Keynote presentations

7/2014	“Networks, Language, Organizations,” Academy of Management, Philadelphia, PA
3/2014	“Analyzing and Simulating Dynamic Networks,” Rand Conference, Arlington, VA
11/2013	“Network Analysis and Visualization,” Approaches to Dynamic Network and Scientometric Analysis within the IC, Washington DC
9/2013	“Open Source Exploitation for Understanding Covert Networks,” Dark Networks, Westpoint, NY
8/2013	“Networks and Agents: The Value of a Multi-Level Approach to Agent-Based Dynamic-Network Modeling.” Statistical and Applied Mathematical Sciences Institute (SAMSI), Raleigh-Durham, NC
5/2013	“Crisis Mapping: Big Data from a Dynamic Network Analytic Perspective,” World Summit on Big Data and Organization Design, Paris, Fr.
10/2012	“Geo-Temporal Dynamic Network Analysis: Lessons Learned from Arab Spring,” GEOINT 2012: Creating the Innovation Advantage, Orlando, FL
9/2012	“Geo-Spatial Network Analysis: Applications and Challenges,” UBICOM, Keynote, Pittsburgh, PA
6/2012	“Political Evolution and Revolution: A Network Assessment of Sudan and Arab Spring Power Transformations,” Keynote, St. Petersburg, Russia
4/2012	“Secondary Actors and Communicative Reach,” Graph Exploitation, Keynote, Boston, MA
4/2012	“Dynamic Network Analysis,” SBP12 International Conference, College Park, MD
3/2012	“Dynamic Network Analysis of the Arab Spring,” IMA, University of Minnesota, Minneapolis, MN